

Imperfect Secrecy in Wiretap Channel II

Fan Cheng[†], Raymond W. Yeung[†], and Kenneth W. Shum[‡]

[†]Department of Information Engineering, The Chinese University of Hong Kong, N.T., Hong Kong

[‡]Institute of Network Coding, The Chinese University of Hong Kong, N.T., Hong Kong
{chengfan, whyeung}@ie.cuhk.edu.hk, wkshum@inc.cuhk.edu.hk

Abstract—In a point-to-point communication system which consists of a sender s , a receiver t and a set of noiseless channels, the sender s wants to transmit a private message to the receiver t through the channels which may be eavesdropped by a wiretapper. The wiretapper can access any one but not more than one set of channels. It is assumed that from each wiretap set, the wiretapper can obtain some partial information about the private message which is measured by the wiretapper's equivocation. The security strategy is to encode the message with some random key. Under these settings, we define an achievable rate tuple in terms of the message, the key and the wiretapper's equivocation and prove a tight rate region of the rate tuples.

Index terms. *Imperfect secrecy, wiretap channel, secret sharing.*

I. INTRODUCTION

Information-theoretic security was launched by Shannon in his seminal paper [7], in which a sender wants to transmit a private message to a receiver with the existence of a wiretapper. This model, referred to as the *Shannon cipher system*, requires that the wiretapper can obtain no information of the message. In this paper, we will refer to it as *perfect security* for ease of discussion. In order to protect the message, the sender encodes the message with a random key which is shared with the receiver a prior but unknown to the wiretapper. The sender transmits the encrypted message in a public channel to the receiver and the receiver can recover the message by the key and encrypted message. For the wiretapper, it can still obtain no information about the private message without the key. The conclusion, known as the *perfect secrecy theorem*, is that the size of the key should be not less than the size of the message if perfect security is required. A recent result by Ho *et al.* in [4] showed an even stronger bound: in the Shannon cipher system, the size of the key is lower bounded by the logarithm of the cardinality of the message alphabet.

The Shannon cipher system was generalized to *secret sharing* by Blakley [1] and Shamir [6]. Ozarow and Wyner [5] also studied a similar problem which they called the *wiretap channel II*. In this model, information is sent to the receiver through a set of point-to-point channels. It is assumed that the wiretapper can access any one but not more than one set of channels, called a wiretap set, out of a collection \mathcal{A} of all possible wiretap sets, where \mathcal{A} is specified by the problem under consideration. In [5], \mathcal{A} consists of all the subsets of the channel set with size r . The strategy to protect the private message is the same as that in the Shannon cipher system. Specifically, they proved a lower bound on the size of the key

which can be attained by a group code. This result is further generalized in Cheng and Yeung [3] for an arbitrary \mathcal{A} . They proved a lower bound on the size of the key and showed that it can be achieved by a linear code.

Imperfect secrecy was independently studied in Yamamoto [9] and Yeung [10] (p. 116). The communication model in [10] is the same as the model described in the Shannon cipher system, except that the wiretapper can obtain a partial information about the message, which is measured by the mutual information between the message and the symbols obtained by the wiretapper. The *imperfect secrecy theorem* states that this mutual information is lower bounded by the difference between the size of the message and the size of the key. In [9], Yamamoto studied source coding problems for Shannon cipher system with correlated source outputs (X, Y) by considering several situations such that both X and Y , only X , or only Y must be transmitted to the receiver, or both X and Y , only X , or only Y must be kept secret from the wiretapper. The admissible region of the cryptogram rate and the key rate for a given security level is derived for each case. A result equivalent to the perfect secrecy theorem is proved in the converse part. When imperfect security is considered in a wiretap network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where \mathcal{V} is the set of nodes and \mathcal{E} is the set of channels, Cai and Yeung in [2] proved two tight bounds on the minimum length of the key and the maximum length of the message, provided that the collection \mathcal{A} of all the possible wiretap sets consists of all the subsets of \mathcal{E} with size r and the information leakage about the message in each wiretap set is at most $i \log q$, where i is a fixed integer satisfying $0 \leq i \leq r$ and q is the size of the alphabet.

Xu and Chen in [8] studied how to communicate securely over a network in which each channel may be noisy or noiseless. Their model is a single-source single-sink acyclic planar network without network coding and the communication between the source and the sink is subject to non-cooperative eavesdropping on each link, namely \mathcal{A} consists of all the subsets of the channel set with a single channel. From each wiretap set in \mathcal{A} , the wiretapper can obtain partial information about the message, which is measured by the wiretapper's equivocation. They defined an achievable rate tuple including the message rate, the key rate and the equivocation rate for each wiretap set. They proved sufficient conditions in terms of the communication rates and the network parameters for provably secure communication, along with an intuitive and efficient coding scheme. Furthermore, the derived achievable rate region is tight for several special cases. In the following,

we refer to this model as the *non-cooperative imperfect secrecy system*.

In this work, we define a security model which generalizes the model in [5]. The communication model is the same as that in [5]. The main difference here is that in our model \mathcal{A} is arbitrary and from each wiretap set in \mathcal{A} , the wiretapper can obtain some information about the message. On the other hand, our model subsumes the noiseless part of the model in [8], since the communication in a single-source single-sink network without network coding can be simplified as a point-to-point system. We also define an achievable rate tuple similar to that in [8] and a tight rate region is proved under these settings.

The rest of the paper is organized as follows. First, we give the problem formulation and introduce some related results in Section II. Then we present our main result on the rate region in Section III including the converse and the achievability. At last, we conclude in Section IV.

II. PROBLEM FORMULATION AND RELATED RESULT

A. Problem Formulation

The communication model in our problem is described as follows:

- The communication is between a transmitter s and a receiver t , which are connected by a set of point-to-point noiseless channels. Let $\mathcal{E} = \{e_1, e_2, \dots, e_h\}$ be the set of channels and $h = |\mathcal{E}|$. For each channel e_i , $1 \leq i \leq h$, the channel capacity is $C_i \log q$, where C_i is an integer. Hence, C_i symbols from a common alphabet \mathcal{F} are transmitted on e_i each time. Denote the symbols transmitted on e_i by Y_{e_i} and $q = |\mathcal{F}|$.
- The message M is generated at the transmitter s according to a uniform distribution on the message set \mathcal{M} . The key K , also generated at the transmitter s , takes value in an alphabet \mathcal{K} according to the uniform distribution, and is independent of M . The transmitter needs to send the encrypted message to the receiver and the receiver needs to recover both the message and the key. The rates of the message and the key are defined as follows.

$$R_M = \frac{H(M)}{\log q}; \quad (1)$$

$$R_K = \frac{H(K)}{\log q}. \quad (2)$$

- Let \mathcal{A} be the set of wiretap sets and $d = |\mathcal{A}|$. For the wiretapper, it can access at most one wiretap set in \mathcal{A} .
- For each wiretap set I_i , $1 \leq i \leq d$, let Y_{I_i} be the symbols transmitted in I_i . It is required that the wiretapper's equivocation $H(M|Y_{I_i})$ is lower bounded by a given constant $R_i \log q$, namely

$$R_i \leq \frac{H(M|Y_{I_i})}{\log q}. \quad (3)$$

The achievable rate tuple is defined as follows.

Definition 1. The encoder is a function f such that $f : \mathcal{M} \times \mathcal{K} \rightarrow \prod_{i=1}^h \mathcal{F}^{C_i}$. The decoder is a function g such that $g : \prod_{i=1}^h \mathcal{F}^{C_i} \rightarrow \mathcal{M} \times \mathcal{K}$. The corresponding rate tuple $(R_M, R_K, R_{i:1 \leq i \leq d})$ is an achievable rate tuple if $g \circ f$ is the identity function and (3) holds for all $i = 1, 2, \dots, d$.

The rate region \mathcal{R} is defined as the set of all achievable rate tuples $(R_M, R_K, R_{i:1 \leq i \leq d})$. In the sequel, we refer to this model as a *cooperative imperfect secrecy system*.

Next, we define the achievable rate tuple by a block code in terms of M , K and Y_{I_i} , $1 \leq i \leq d$.

Definition 2. A rate tuple of $(R_M, R_K, R_{i:1 \leq i \leq d})$ is achievable by block codes if there exists a sequence of (M_n, K_n) such that

$$R_M = \lim_{n \rightarrow \infty} \frac{1}{n} \frac{\log |\mathcal{M}_n|}{\log q}; \quad (4)$$

$$R_K = \lim_{n \rightarrow \infty} \frac{1}{n} \frac{\log |\mathcal{K}_n|}{\log q}; \quad (5)$$

$$R_i \leq \lim_{n \rightarrow \infty} \inf \frac{1}{n} \frac{H(M_n|Y_{I_i,n})}{\log q}, 1 \leq i \leq d; \quad (6)$$

where $M_n \in \mathcal{M}_n \subseteq \mathcal{M}^n$, $K_n \in \mathcal{K}_n \subseteq \mathcal{K}^n$, and $Y_{I_i,n} \in \mathcal{F}^n$.

In the sequel, we assume that the base of the logarithm in the entropy quantities (e.g., $H(X), I(X; Y)$) is q . Then the factor $\frac{1}{\log q}$ can be omitted in (1)-(6).

B. Related Result

1) *Perfect and Imperfect Secrecy*: The perfect secrecy theorem in [7] is stated as follows.

Theorem 1 (Perfect Secrecy Theorem). Let X be the plain text, Y be the cipher text, and K be the key in a secret key cryptosystem. If perfect secrecy is achieved, i.e., $I(X; Y) = 0$, then

$$H(K) \geq H(X). \quad (7)$$

In the wiretap network model [2], the following result similar to the perfect secrecy theorem was proved.

Theorem 2.¹ In a wiretap network, let K be the key and Y_I be the symbols transmitted in wiretap set I . Then

$$H(K) \geq H(Y_I). \quad (8)$$

As a generalization of the perfect secrecy theorem, the imperfect secrecy theorem in [10] (p. 116) is stated below.

Theorem 3 (Imperfect Secrecy Theorem). Let X be the plain text, Y be the cipher text, and K be the key in a secret key cryptosystem. Then

$$I(X; Y) \geq H(X) - H(K). \quad (9)$$

In the above theorem, if $I(X; Y) = 0$, then (9) becomes (7), i.e., the theorem reduces to the perfect secrecy theorem. In [9], it was proved that for any secret key cryptosystem,

$$H(K) \geq H(X|Y), \quad (10)$$

which is equivalent to (9).

¹This theorem can be found in Appendix A, Equation (27) of [2].

2) *Secure Coding over Networks*: The system model in [8] is a single-source single-sink directed acyclic network with the assumption that each wiretapper can access only one channel and there is no network coding in the network. Each channel in the network may be noisy or noiseless.

When all the channels in the network are noiseless, the network can be simplified as a point-to-point communication system, in which each channel is a path from the source node to the destination node in the original network and the set of wiretap sets \mathcal{A} is arbitrary. Hence our model subsumes the non-cooperative model for this special case.

In [8], an achievable rate region of rate tuples was obtained for noisy channels, and the region was shown to be tight for several special cases. Based on the achievable rate region, they also gave an algorithm for constructing a secure code on the network.

The achievable rate region for noiseless channels is stated below.

Theorem 4 (Theorem 2, [8]). *A rate tuple (R_M, R_K, R_e) , $e \in \mathcal{E}$, is achievable, if there exist auxiliary numbers r_e such that*

$$\begin{aligned} 0 &\leq r_e \leq R_M + R_K; \\ 0 &\leq R_e \leq R_M; \\ 0 &\leq R_M + R_K \leq \min_{Cut} \sum_{e \in \mathcal{E}_{cut}} r_e; \\ r_e &\leq C_e; \\ R_e &\leq R_M + R_K - r_e. \end{aligned}$$

In the above, R_e and C_e correspond to R_i and C_i in our formulation; \mathcal{E}_{Cut} is the set of channels across a given cut Cut .

III. RATE REGION OF THE RATE TUPLE

The main result of this paper is a characterization of the rate region \mathcal{R} given by the following theorem.

Theorem 5. *A rate tuple $(R_M, R_K, R_{i:1 \leq i \leq d})$ is in \mathcal{R} if and only if*

$$R_M = \sum_{i=1}^h r_i - R_K; \quad (11)$$

$$R_M \geq R_i, \quad 1 \leq i \leq d; \quad (12)$$

$$R_i \geq 0, \quad 1 \leq i \leq d; \quad (13)$$

$$R_K \geq 0; \quad (14)$$

where r_i 's satisfy

$$0 \leq r_i \leq C_i, \quad 1 \leq i \leq h; \quad (15)$$

$$\sum_{e_i \in I_j} r_i \leq R_K + R_M - R_j, \quad 1 \leq j \leq d. \quad (16)$$

A. Converse

In this section, we prove that if $(R_M, R_K, R_{i:1 \leq i \leq d}) \in \mathcal{R}$, then the constraints (11)-(16) hold. The constraints (13) and (14) are obvious.

We first prove the constraint (12). By the constraint (3),

$$R_i \leq H(M|Y_{I_i}) \leq H(M) = R_M. \quad (17)$$

Hence the constraints (12)-(14) hold.

Let's consider an equivalent condition of constraint (3). For all $1 \leq i \leq d$, let

$$c_i = R_M - R_i = H(M) - R_i. \quad (18)$$

The constraint (3) is equivalent to

$$I(Y_{I_i}; M) \leq H(M) - R_i.$$

Namely

$$0 \leq I(Y_{I_i}; M) \leq c_i. \quad (19)$$

By (17) and (18),

$$0 \leq c_i \leq R_M.$$

Next, we prove a lemma which generalizes Theorem 2.

Lemma 1. *In a cooperative imperfect secrecy system, let M be the message, K be the key and Y_I be the symbols transmitted in wiretap set I . Then*

$$I(Y_I; M) \geq H(Y_I) - H(K). \quad (20)$$

Proof: Since $I(M; K) = 0$ and $H(Y_I|M, K) = 0$,

$$\begin{aligned} I(Y_I; M) &= H(Y_I) - H(Y_I|M) \\ &\geq H(Y_I) - H(M, K|M) \\ &= H(Y_I) - H(K|M) \\ &= H(Y_I) - H(K). \end{aligned}$$

■

In the next theorem, we prove the constraints (11), (15), and (16).

Lemma 2. *Any tuple $(R_M, R_K, R_{i:1 \leq i \leq d}) \in \mathcal{R}$ satisfies*

$$R_M = \sum_{i=1}^h r_i - R_K,$$

where r_i 's satisfy

$$0 \leq r_i \leq C_i, \quad 1 \leq i \leq h;$$

$$\sum_{e_i \in I_j} r_i \leq R_K + R_M - R_j, \quad 1 \leq j \leq d.$$

Proof: By Lemma 1 and the inequality (19), for each wiretap set I_i ,

$$H(Y_{I_i}) - H(K) \leq I(Y_{I_i}; M) \leq c_i.$$

Namely,

$$H(Y_{I_i}) \leq H(K) + c_i = R_K + c_i.$$

For each channel e_i , $1 \leq i \leq h$,

$$H(Y_{e_i}) \leq C_i.$$

Since $Y_{(e_i:1 \leq i \leq h)}$ is a function of (M, K) and (M, K) can be recovered by $Y_{(e_i:1 \leq i \leq h)}$,

$$H(Y_{(e_i:1 \leq i \leq h)}) = H(M, K) = H(M) + H(K). \quad (21)$$

Hence,

$$H(M) = H(Y_{(e_i:1 \leq i \leq h)}) - H(K),$$

which is equivalent to

$$R_M = H(Y_{(e_i:1 \leq i \leq h)}) - R_K.$$

For $1 \leq i \leq h$, let

$$r_i = H(Y_{e_i} | Y_{(e_1, e_2, \dots, e_{i-1})}).$$

Then for all I_j , $1 \leq j \leq d$,

$$r_i \leq H(Y_{e_i} | Y_{(e_l: e_l \in I_j, l < i)}).$$

Furthermore

$$\begin{aligned} R_M &= H(Y_{(e_i:1 \leq i \leq h)}) - R_K \\ &= \sum_{i=1}^h H(Y_{e_i} | Y_{(e_1, e_2, \dots, e_{i-1})}) - R_K \\ &= \sum_{i=1}^h r_i - R_K; \\ 0 &\leq r_i \leq H(Y_{e_i}) \leq C_i; \\ \sum_{e_i \in I_j} r_i &\leq \sum_{e_i \in I_j} H(Y_{e_i} | Y_{(e_l: e_l \in I_j, l < i)}) = H(Y_{I_j}) \\ &\leq R_K + c_j = R_K + R_M - R_j, 1 \leq j \leq d, \end{aligned}$$

which completes the proof. \blacksquare

B. Achievability

In this section, we prove that $(R_M, R_K, R_{i:1 \leq i \leq d}) \in \mathcal{R}$ if there exists (r_1, r_2, \dots, r_h) such that the constraints (11)-(16) are satisfied.

In the following, a special code in which the symbols sent on the channels are mutually independent is studied. We design a block code with length n as follows. The sender generates M and K at rates R_M and R_K , respectively, and sends symbols on each channel e_i ($1 \leq i \leq h$) at rate r_i . Next, we prove that the tuple $(R_M, R_K, R_{i:1 \leq i \leq d})$ can be attained by a linear code.

Let the symbols on channel e_i ($1 \leq i \leq h$) be X_i . $\lfloor x \rfloor$ is the floor function. Let

$$c'_i = \lfloor nc_i \rfloor; \quad (22)$$

$$C'_i = \lfloor nC_i \rfloor; \quad (23)$$

$$n_M = \lfloor nR_M \rfloor = \lfloor nH(M) \rfloor; \quad (24)$$

$$n_K = \lfloor nR_K \rfloor = \lfloor nH(K) \rfloor; \quad (25)$$

$$n_i = \lfloor nr_i \rfloor = \lfloor nH(X_i) \rfloor, 1 \leq i \leq h. \quad (26)$$

Thus, by (11), (15), and (16), n_M , n_K , and (n_1, n_2, \dots, n_h) satisfy that

$$n_M = \sum_{i=1}^h n_i - n_K; \quad (27)$$

$$0 \leq n_i \leq C'_i, 1 \leq i \leq h; \quad (28)$$

$$\sum_{e_j \in I_i} n_j \leq n_K + c'_i, 1 \leq i \leq d. \quad (29)$$

Usually, there may be rounding errors in (27)-(29). Since real numbers can be approximated by rational numbers, we can assume that the variables in (22)-(26), i.e., c_i, R_M, R_K , and $r_i, 1 \leq i \leq h$, are rational numbers. There exist infinitely many n such that nc_i, nC_i, nR_M, nR_K , and $nr_i, 1 \leq i \leq h$, are integer numbers. Hence, the rounding errors can be omitted here.

When $n \rightarrow \infty$,

$$\left(\frac{n_1}{n}, \frac{n_2}{n}, \dots, \frac{n_h}{n}\right) \rightarrow (r_1, r_2, \dots, r_h);$$

$$\sum_{i=1}^h \frac{n_i}{n} \rightarrow R_M + R_K.$$

Hence, when $n \rightarrow \infty$, $(\frac{n_1}{n}, \frac{n_2}{n}, \dots, \frac{n_h}{n})$ and (r_1, r_2, \dots, r_h) are equivalent.

For a matrix A , we write the number of rows and columns of A as $\text{row}(A)$ and $\text{col}(A)$, respectively. The following two lemmas are instrumental in the subsequent proofs.

Lemma 3. Let F_q be a finite field of size q , A, B be given matrices and (A, B) be the concatenated matrix of A and B . Let $Y = AM + BK$, where $\text{rank}(A, B) = \text{row}(A, B)$. If M and K are uniformly distributed on F_q^m and F_q^k , respectively, and $I(M; K) = 0$, then

$$I(Y; M) = \text{rank}(A, B) - \text{rank}(B).$$

Proof:

$$\begin{aligned} I(Y; M) &= H(Y) - H(Y|M) \\ &= H(Y) - H(AM + BK|M) \\ &= H(Y) - H(BK|M) \\ &= H(Y) - H(BK) \\ &= \text{rank}(A, B) - \text{rank}(B). \end{aligned}$$

\blacksquare

Lemma 4 (Lemma 3, [2]). Let V_1, V_2, \dots, V_m be vector subspaces in F_q^n , and $\dim(V_i) = d_i$ ($1 \leq i \leq m$). If $d \geq 0$ and $d + d_i \leq n$ ($1 \leq i \leq m$), then for $q > m$, there exists a vector subspace V of F_q^n , such that $\dim(V) = d$ and $\dim(V \oplus V_i) = \dim(V) + \dim(V_i)$ ($1 \leq i \leq m$).

The remaining of this paper is largely about the following theorem.

Theorem 6. When $q > |\mathcal{A}|$ is a prime power, if the integer tuple (n_1, n_2, \dots, n_h) satisfies (27)-(29), then there exists a linear code such that $H(M^n) = n_M$ and $H(K^n) = n_K$.

Proof: The code can be constructed as follows:

Let the finite field F_q be the common alphabet of M and K and the common alphabet of all the channels $\mathcal{F} = F_q$. The symbols transmitted on channel e_i ($1 \leq i \leq h$) is taken from $F_q^{n_i}$, which means there are n_i symbols from F_q transmitted on e_i . Let $x_1, x_2, \dots, x_{n_M+n_K}$ be all the symbols to send, where the first n_1 symbols are sent on e_1 , the next n_2 symbols are sent on e_2, \dots , and the last n_h symbols are sent on e_h . We construct x_i 's according to their positions in the sequence.

Generate n_K mutually independent symbols $K = (k_1, k_2, \dots, k_{n_K})$ from F_q . Transmit K at the first n_K positions, i.e., $1, 2, \dots, n_K$. Namely,

$$x_i = k_i, 1 \leq i \leq n_K.$$

Generate $n_M (= \sum_{i=1}^h n_i - n_K)$ mutually independent message symbols $(m_1, m_2, \dots, m_{n_M})$ from F_q . For the remaining n_M positions in $\{e_i : 1 \leq i \leq h\}$, transmit the encrypted message by the encoding function:

$$x_i = m_{i-n_K} + b_i K, \quad n_K + 1 \leq i \leq n_K + n_M, \quad (30)$$

where $b_i \in F_q^{n_K}$ is a row vector to be determined in the following steps.

We need to construct $\{b_i : n_K + 1 \leq i \leq n_K + n_M\}$ such that:

- (a) These $n_K + n_M$ symbols are mutually independent, which is equivalent to that both M and K can be recovered at node t .

From the previous discussions, we can see that receiver t can recover K from the symbols in the first n_K positions and by equation (30), M can be also recovered by

$$m_{i-n_K} = x_i - b_i K, \quad n_K + 1 \leq i \leq n_K + n_M.$$

Hence, the required condition is satisfied.

- (b) The constraint (19) (which is equivalent to (3)) should hold for all the wiretap sets.

For $\{x_1, x_2, \dots, x_{n_M+n_K}\}$,

$$\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_{n_M+n_K} \end{pmatrix} = (A \mid B) \begin{pmatrix} M^n \\ K^n \end{pmatrix},$$

where

$$(A \mid B) = \left(\begin{array}{c|c} \mathbf{0} & I_{n_K \times n_K} \\ I_{n_M \times n_M} & \begin{matrix} b_{n_K+1} \\ \dots \\ b_{n_K+n_M} \end{matrix} \end{array} \right).$$

In the above, $\mathbf{0}$ is an $n_K \times n_M$ zero matrix and $I_{n_K \times n_K}$ is an $n_K \times n_K$ identity matrix. Recall that the symbols obtained in wiretap set $I_i = \{e_{i_1}, e_{i_2}, \dots, e_{i_{|I_i|}}\}$ are Y_{I_i} , $1 \leq i \leq d$. Then

$$Y_{I_i} = \begin{pmatrix} x_{i_1} \\ x_{i_2} \\ \dots \\ x_{i_{|I_i|}} \end{pmatrix} = (A_{I_i}, B_{I_i}) \begin{pmatrix} M^n \\ K^n \end{pmatrix},$$

where A_{I_i} and B_{I_i} are the corresponding sub-matrices of A and B , respectively. Consequently, we turn to a sufficient condition of b_i 's for the constraint (19). Since $x_1, x_2, \dots, x_{n_M+n_K}$ are mutually independent,

$$\text{rank}(A_{I_i}, B_{I_i}) = \text{row}(A_{I_i}, B_{I_i}) = \sum_{e_j \in I_i} n_j. \quad (31)$$

By Lemma 3,

$$\begin{aligned} I(Y_{I_i}; M) &= \text{rank}(A_{I_i}, B_{I_i}) - \text{rank}(B_{I_i}) \\ &= \sum_{e_j \in I_i} n_j - \text{rank}(B_{I_i}). \end{aligned}$$

If the constraint (19) holds, then

$$I(Y_{I_i}; M) \leq n \times c_i = c'_i.$$

Hence B_{I_i} should satisfy that

$$\sum_{e_j \in I_i} n_j - \text{rank}(B_{I_i}) \leq c'_i.$$

Namely,

$$\text{rank}(B_{I_i}) \geq \sum_{e_j \in I_i} n_j - c'_i, \quad \text{for all } 1 \leq i \leq d. \quad (32)$$

For $\sum_{e_j \in I_i} n_j$, by condition (29), since

$$\sum_{e_j \in I_i} n_j \leq n_K + c'_i,$$

we obtain that

$$\sum_{e_j \in I_i} n_j - c'_i \leq n_K = \text{col}(B_{I_i}). \quad (33)$$

By (31),

$$\begin{aligned} \sum_{e_j \in I_i} n_j - c'_i &= \text{row}(A_{I_i}, B_{I_i}) - c'_i \\ &= \text{row}(B_{I_i}) - c'_i \leq \text{row}(B_{I_i}). \end{aligned} \quad (34)$$

In summary, by (32)-(34), it is required to construct b_i 's such that

$$\text{rank}(B_{I_i}) \geq \sum_{e_j \in I_i} n_j - c'_i, \quad 1 \leq i \leq d, \quad (35)$$

where

$$\begin{aligned} \sum_{e_j \in I_i} n_j - c'_i &\leq \text{col}(B_{I_i}) = n_K; \\ \sum_{e_j \in I_i} n_j - c'_i &\leq \text{row}(B_{I_i}) = \sum_{e_j \in I_i} n_j. \end{aligned}$$

For (35), it suffices to construct b_i 's such that for all i , $1 \leq i \leq d$,

$$\begin{aligned} \text{rank}(B_{I_i}) &= \min\{\text{row}(B_{I_i}), \text{col}(B_{I_i})\} \\ &= \min\left\{\sum_{e_j \in I_i} n_j, n_K\right\}. \end{aligned} \quad (36)$$

Next, we construct b_i 's by mathematical induction.

Initially, for $1 \leq i \leq n_K$, $b_i = (\underbrace{0, 0, \dots, 0}_{i-1}, 1, 0, \dots, 0)$. Let

$$Y_{I_i}^j = \begin{pmatrix} x_{i_1} \\ x_{i_2} \\ \dots \\ x_{i_l} \end{pmatrix} = (A_{I_i}^j, B_{I_i}^j) \begin{pmatrix} M \\ K \end{pmatrix},$$

where x_{i_l} 's are the symbols sent in I_i with $1 \leq i_l \leq j$. Thus, $Y_{I_i}^j$ is a sub-vector of Y_{I_i} up to index j and so are $A_{I_i}^j$ and $B_{I_i}^j$. When $j = n_M + n_K$, $Y_{I_i}^j = Y_{I_i}$, $A_{I_i}^j = A_{I_i}$ and $B_{I_i}^j = B_{I_i}$. Since $B_{I_i}^{n_K}$ is a sub-matrix of the $n_K \times n_K$ identity matrix $I_{n_K \times n_K}$,

$$\text{rank}(B_{I_i}^{n_K}) = \text{row}(B_{I_i}^{n_K}),$$

which means inequality (36) holds.

Suppose that when $k = l \geq n_K$, $\{b_i : 1 \leq i \leq l\}$ have been constructed successfully. Thus for all i , $1 \leq i \leq d$,

$$\begin{aligned} \text{rank}(B_{I_i}^l) &= \min\{\text{row}(B_{I_i}^l), \text{col}(B_{I_i}^l)\} \\ &\leq \min\left\{\sum_{e_j \in I_i} n_j, n_K\right\}. \end{aligned}$$

For b_{l+1} , it is required that for all wiretap set I_i ($1 \leq i \leq d$) accesses x_{l+1} , if $\text{rank}(B_{I_i}^l) < \min\left\{\sum_{e_j \in I_i} n_j, n_K\right\}$,

then

$$\text{rank}(B_{I_i}^{l+1}) = \text{rank}(B_{I_i}^l) + 1,$$

where $B_{I_i}^{l+1}$ and $B_{I_i}^l$ satisfy

$$B_{I_i}^{l+1} = \begin{pmatrix} B_{I_i}^l \\ b_{l+1} \end{pmatrix}.$$

The existence of b_{l+1} can be guaranteed by Lemma 4. Hence b_{l+1} has been constructed successfully. In the above proof, by Lemma 4, $q > |\mathcal{A}| = d$ is sufficient for the existence of b_i 's.

By mathematical induction, we complete the proof.

Hence, b_i 's are successfully constructed, which completes the proof. \blacksquare

IV. CONCLUSION

In this paper, we have proved a tight rate region of the rate tuples in the cooperative imperfect secrecy model by a linear program, in which the key idea is from the imperfect secrecy theorem. The result can be treated as a bound from a cut-set if network coding is allowed in a general wiretap network. Although for a general case, the rate region of rate tuples is still open, our result has paved the way for the further discussion on this problem.

ACKNOWLEDGMENT

This work was partially supported by a grant from the University Grants Committee (Project No. AoE/E-02/08) of the Hong Kong Special Administrative Region, China.

REFERENCES

- [1] G. R. Blakley, "Safeguarding cryptographic keys," in *Proceedings of the National Computer Conference*, 48: 313- 317, 1979.
- [2] N. Cai and R. W. Yeung, "Secure Network Coding on a Wiretap Network," *IEEE Trans. on Inform. Theory*, 57(1):424-435, Jan. 2011.
- [3] F. Cheng and R. W. Yeung, "Performance Bounds in Secure Network Coding," *IEEE International Symposium on Network Coding (NetCod)*, Jul. 2011.
- [4] S.-W. Ho, T. Chan, and C. Uduwerelle, "Error-free perfect-secrecy systems," *IEEE International Symposium on Information Theory (ISIT)*, pp. 1613-1617, Aug. 2011.
- [5] L. H. Ozarow and A. D. Wyner, "Wire-tap Channel II," *AT&T Bell Labs. Tech. J.*, 63: 2135-2157, 1984.
- [6] A. Shamir, "How to share a secret," *Comm. ACM*, 22: 612- 613, 1979.
- [7] C. E. Shannon, "Communication theory of secrecy systems," *Bell Sys. Tech. Journal* 28, pp. 656-715, 1949.
- [8] J. Xu and B. Chen, "Secure coding over networks," *IEEE International Symposium on Information Theory (ISIT)*, pp. 2116-2120, Jul. 2009.
- [9] H. Yamamoto, "Coding theorems for Shannon's cipher system with correlated source outputs, and common information," *IEEE Trans. Inform. Theory*, vol. 40, no. 1, pp. 85-95, 1994.
- [10] R. W. Yeung, *A First Course in Information Theory*, Kluwer Academic/Plenum Publishers, 2002.